Renewable and Sustainable Energy Reviews 15 (2011) 1556–1566

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Influence of the input layer signals of ANNs on wind power estimation for a target site: A case study

Sergio Velázquez[a], José A. Carta[b,*], J.M. Matías[c]

[a] Department of Electronics and Automatics Engineering, Universidad de Las Palmas de Gran Canaria, Campus de Tafira s/n, 35017 Las Palmas de Gran Canaria, Canary Islands, Spain
[b] Department of Mechanical Engineering, Universidad de Las Palmas de Gran Canaria, Campus de Tafira s/n, 35017 Las Palmas de Gran Canaria, Canary Islands, Spain
[c] Department of Statistics, University of Vigo, Lagoas Marcosende, 36200 Vigo, Spain

ARTICLE INFO

Article history:
Received 20 August 2010
Accepted 12 November 2010
Available online 12 January 2011

Keywords:
Artificial Neural Networks
Wind power estimation
Wind direction
Multilayer perceptron
Backpropagation algorithm
Mean absolute relative error

ABSTRACT

In the scientific literature concerning renewable energies there have been proposals for the use of Artificial Neural Networks (ANNs) as a tool to estimate the mean wind speed at a target station for which only incomplete data series are available. In general terms, the wind speeds recorded at neighbouring reference stations are used as signals of the input layer of multilayer perceptron (MLP) architectures.

An analysis is undertaken in this paper of the extent to which estimation of the wind speed, the wind power density and the power output of an installed wind turbine at a target site is affected by the number of reference stations, the degree of correlation between the wind speeds of the reference stations and the target station, the wind direction and the manner in which the direction signal is introduced into the input layer.

For the purposes of this study mean hourly wind speeds and directions were used which had been recorded at twenty two weather stations located in the Canary Archipelago (Spain) in 2002 and 2006. The power-wind speed characteristic curves of five wind turbines of different rated power were also used.

The MLP architectures were trained using the backpropagation algorithm. Amongst other conclusions reached with the tests performed with the set of data not used in the training nor in the validation of the models is that the estimation errors tend to decrease as the number of reference stations used increases, independently of the existing correlation coefficient values between the reference stations and the candidate station. It was also observed that if input signals of angular wind direction and wind speed are used then the number of reference stations required to achieve a certain decrease in the error is lower than the number required if only wind speed input signals are used. By using multiple reference stations with input signals of angular wind direction and speed, error reductions have been achieved for some stations in the estimation of the wind power density and the power output of wind turbines, with respect to the case of a single reference station and identical input signals, of 75% and 62%, respectively.
© 2010 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author. Tel.: +34 928 45 14 83; fax: +34 928 45 14 84.
E-mail address: jcarta@dim.ulpgc.es (J.A. Carta).

1364-0321/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rser.2010.11.036

Nomenclature

AEMET	State Meteorological Agency (Spanish initials: AEMET) of the Ministry of the Environment and Rural and Marine Environs of the Spanish Government
ANN	Artificial Neural Network
BN	Bayesian network
CC	correlation coefficient between the estimated wind speeds and the measured wind speeds of the candidate station. Eq. (2)
ITC	Instituto Tecnológico de Canarias (Canary Islands Technological Institute)
k	number of neurons of the hidden layer of an ANN
M	number of nodes in which wind turbine manufacturers discretize their power-wind speed curves
MARE	mean absolute relative error. Eq. (1)
MCP	Measure-Correlate-Predict algorithms
MLP	multilayer perceptron
N	number of reference stations
n	number of data used in the test subset. Eqs. (1) and (2)
O_i	Estimated data. Eqs. (1) and (2)
$\bar{O}$	mean of the estimated O_i data. Eq. (2)
P_i	Wind power density (in W m^{-2}). Eq. (4)
$PWT(v)$	power-wind speed curve of a wind turbine. Eq. (5)
Pwt_i	ordinate of the i node of the power-wind speed curve of a wind turbine (in kW)
R	correlation coefficient measured between the wind speeds of a reference station and those of the candidate station (Tables 2 and 3).
T_i	measured data. Eqs. (1) and (2)
$\bar{T}$	mean of the measured T_i data. Eq. (2)
vc	wind speeds estimated with the ANN models used (in m s^{-1})
$V_{cut\text{-}in}$	cut-in speed of a wind turbine
$V_{cut\text{-}out}$	cut-out wind speed of a wind turbine
vcx, vcy	Cartesian components of the wind velocity of a candidate station
V_i	abscissa of the i node of the power-wind speed curve of a wind turbine (in m s^{-1})
v_i	mean hourly wind speeds (in m s^{-1})
vrx, vry	Cartesian components of the wind velocity of a reference station
WS	weather station
θ	angular direction of the wind speed (°)
ρ_i	air density (in kg m^{-3})
$\bar{\rho}$	mean air density, corresponding to standard conditions (in kg m^{-3})

1. Introduction

Various methods have been published in the scientific literature concerned with renewable energies to estimate the mean wind speed, for a given period of time, at a candidate or target site for which an incomplete series of wind data is available. Some authors have used these methods to estimate missing values in an annual series of wind speeds recorded at a particular site [1]. However, these methods have fundamentally been proposed for estimation of the long-term wind speeds (ten years or more) [2–5] at a candidate site for the exploitation of wind energy for which only short-term data are available (normally of one year). This application is justified by the fact that, as a result of interannual variability, knowledge of the wind performance over just one year is insufficient to estimate its mean performance [6]. This variability has an impact on the energy generated by a wind turbine as it is very sensitive to the wind speed that acts on it over its useful life. So, this variability explains the need to estimate the long-term wind speed for the purpose of assessing the economic feasibility of the investment that needs to be made in the installation of a wind energy conversion system.

To achieve these goals the methods that have been proposed require the use of long-term wind data records from nearby reference weather stations. Fig. 1 shows a schematic diagram of the procedure that is commonly employed. Using the series of wind data recorded at the reference stations and the candidate station for the common short-term period of data availability the architecture of the model is selected, trained, validated and tested [7] (this stage is indicated by the circled number 1 in Fig. 1). The known long-term data series from the reference stations are then fed into the model (this stage is indicated by the circled number 2 in Fig. 1) and the unknown long-term wind speeds and directions for the candidate or target station are estimated (this stage is indicated by the circled number 3 in Fig. 1). It is anticipated that the better the estimation models designed in stage 1, the better the fit will be of the estimations made over the long-term [8].

Amongst the various methods used are the traditional Measure-Correlate-Predict (MCP) algorithms [5,8–10] and methods that employ automatic learning techniques [11–18]. The former normally use a single reference station in the estimation of the unknown data of the target station. With some exceptions [5,10], most traditional MCP methods use linear regression algorithms [19] to characterise the relation between the wind speeds of the candidate site and the reference site. MCP methods generally bin or group together hourly wind speeds according to the wind direction at the reference site. For this purpose, the data are divided into a particular number of direction sectors, normally eight of 45° or twelve of 30°. The parameters which define the regression lines of the model are then estimated for each direction sector [8]. That is, the parameters of the models are functions of the wind speed and direction.

The second type are algorithms which take their inspiration from the world of biology, as is the case of Artificial Neural Networks (ANNs) [11–17] and statistical learning algorithms such as Bayesian networks (BNs) [18]. Automatic learning techniques allow the use of information available from various reference stations and this possibility has been employed in most of the studies published. So, these intelligent models generally make use of regional information about the wind resource when it comes to estimating the wind data for a target station.

Though there are exceptions [15], in the supervised learning of the net the ANN models employed in the scientific literature use as input layer signals meteorological variables recorded at selected reference stations. A high percentage of these models only use as input layer signals the wind speeds recorded at the selected reference stations. The topologies of the ANNs used only have a single neuron in the output layer, with the output signal being the estimated wind speed at the candidate site. That is, most of the models used ignore the influence that the wind direction at the reference stations might have on estimation of the wind speeds at the target station [11–15]. In [16] a study was conducted of the influence of the wind direction, in angular magnitude, on the estimation of wind speed in complex terrain. However, of the five weather stations employed in that study the same target station was always used and the coefficients of correlation between stations were of the same order.

An analysis is undertaken in this paper of the extent to which estimation of the wind speed, the wind power density and the power output of an installed wind turbine at a candidate site is affected by the number of reference stations, the degree of cor-

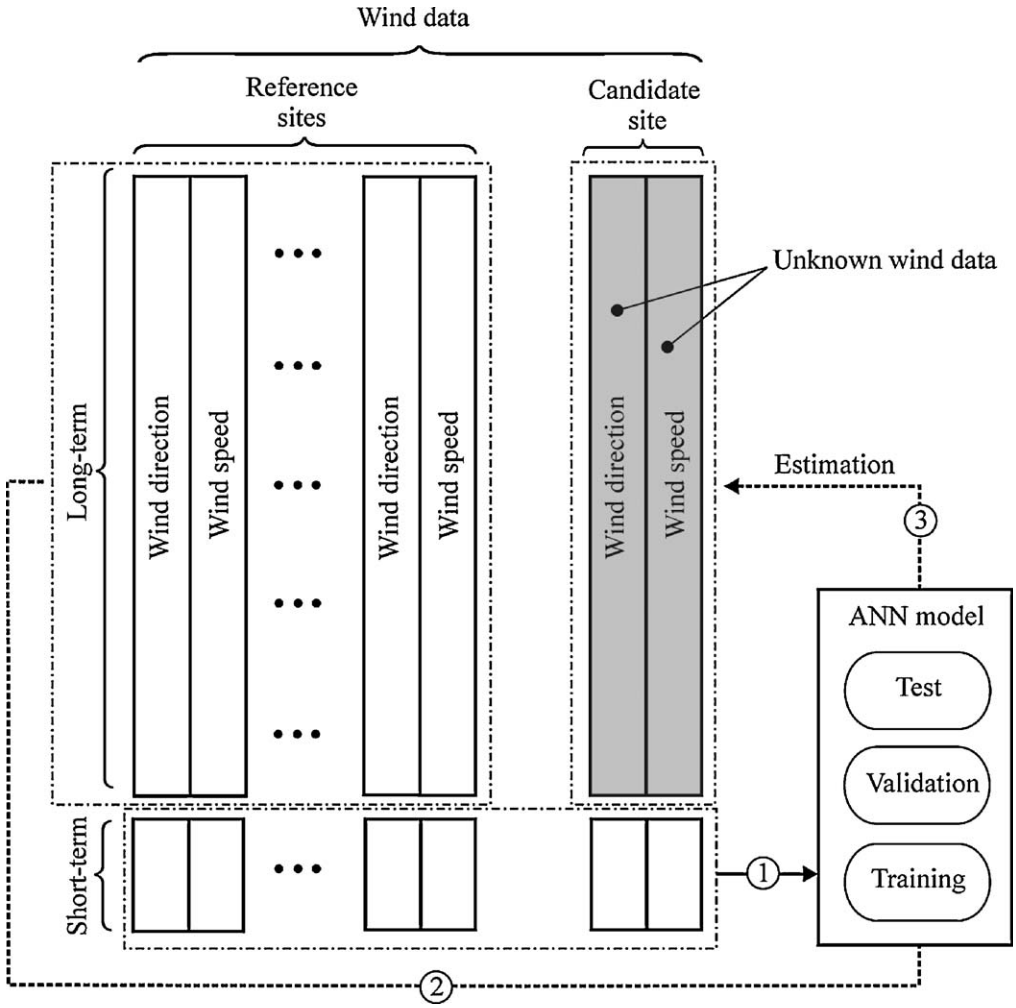


Fig. 1. Stages of the process of estimation of the long-term wind speeds.

relation between the wind speeds of the reference stations and the target station, the wind direction and the manner in which the direction signal is introduced into the input layer. This analysis centres on stage 1, as indicated in Fig. 1. That is, in the stage of training, validation and testing of the models. The conclusions that are obtained from the analysis performed at this stage will enable recommendations to be made when it comes to using ANNs in the long-term estimation of wind speeds and the energy generated by a wind turbine.

For the purpose of this study mean hourly wind speeds and directions recorded at twenty two weather stations installed in six of the seven islands that make up the Canary Archipelago (Spain) were used. The power-wind speed characteristic curves of five wind turbines of different rated power were also used [20,21].

2. Architecture of the ANNs used

The ANNs used in this paper were comprised of three layer networks with feedforward connections. More specifically, multilayer perceptron topologies (MLPs) were used [22,23]. A single hidden layer of neurons was employed so as not to increase the training time. This architecture has demonstrated its ability to approximate satisfactorily any continuous transformation [22,23] and has been proposed by several authors [11,14,16]. Various preliminary tests were carried out to choose the number of hidden neurons, varying the number of input signals. These tests showed that the use of a number of neurons above twenty increased the training time but provided no improvement in the results. For this reason the number of neurons in the hidden layer was set at twenty. The number of neurons of the input layer varies depending on the hypotheses being analysed. That is, it depends on the number of reference stations used and on consideration or otherwise of the wind speed direction signal. In most of the hypotheses that are considered the output layer has a single neuron. The output signal provided by that neuron is the estimated mean hourly wind speed at the target station. In the hypothesis where the ANN input signals, for each reference station, are the components, depending on a system of Cartesian axes, of the wind velocity, the output layer of the network contains two neurons. The outputs provided by these neurons are the components, depending on each of the axes of the Cartesian reference system, of the estimated mean hourly wind velocity at the target station.

The designed architectures were trained using the backpropagation algorithm with a sigmoidal activation function [22,23].

The Levenberg–Marquardt algorithm was used for minimization of the mean square error committed during the learning [19,22].

For training and testing of the network the year long series of data available for the reference and target stations were divided into three random subsets of different data: training data, validation data and test data [7]. The proportion of data selected for each of these processes was 60%, 20% and 20%, respectively.

The training data subset is used for estimation of the weights of the ANNs. The validation data subset is used to check the training progress of the ANNs, optimizing their parameters. That is, they are

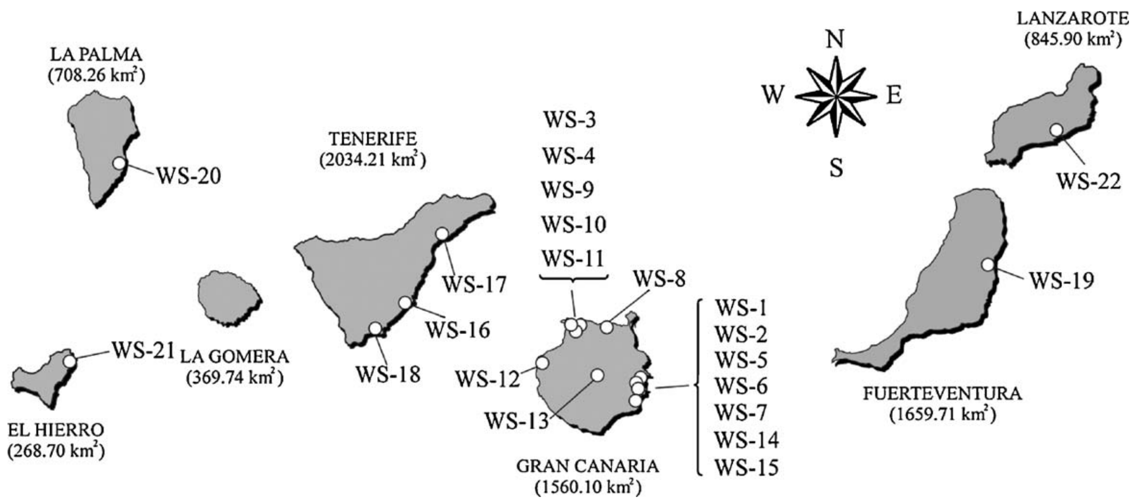


Fig. 2. Location of the weather stations used.

used to measure the degree of generalisation of the ANNs. The test data subset, which were used in neither the training nor validation stages, is used to calculate the error percentage of the final optimized network. That is, it constitutes an independent measure of the performance of the network after training.

The various experiments were carried out using the 'trainlm' algorithm available in the MATLAB Neural Network Toolbox [24].

3. Meteorological data and wind turbines used

The meteorological data used in this paper (mean hourly wind speeds and directions) were recorded at twenty two weather stations installed in six of the seven islands that make up the Canary Archipelago (Spain) (Fig. 2). It can be seen that all the recording devices are located along the coasts of the islands, except for WS-13 which is in the centre of Gran Canaria island. The data series used were provided by the State Meteorological Agency (Spanish initials: AEMET) of the Ministry of the Environment and Rural and Marine Environs of the Spanish Government and by the Canary Islands Technological Institute (Spanish initials: ITC).

The first column of Table 1 shows the codes assigned to each weather station. Also shown in Table 1 are the measuring period, the source of the data, the height above ground level of the measuring devices, the geographic coordinates of the stations (latitude, longitude and altitude), the annual mean wind speed and the standard deviation. Tables 2 and 3 show the linear correlation coefficients between the mean hourly wind speeds of the various stations in 2002 and 2006, respectively. Despite the relatively short distances that lie between the sites of the weather stations, as a result of the orography of the areas where they are located the

Table 1
Weather stations used in the study.

Weather stations	Years	Source	Height a.g.l. (m)	Geographical coordinates			Annual mean (m/s)	Standard deviation (m/s)
				Latitude north	Longitude west	Altitude (m)		
WS-1	2002	AEMET	10	27°55′44″	15°23′20″	16	7.35	3.71
WS-1	2006	AEMET	10	27°55′44″	15°23′20″	16	7.09	3.49
WS-2	2002	ITC	10	27°56′46″	15°23′06″	9	6.11	3.06
WS-3	2002	ITC	10	28°07′30″	15°40′37″	472	7.51	3.88
WS-4	2002	ITC	20	28°07′30″	15°40′37″	472	7.67	3.85
WS-5	2006	ITC	10	27°51′36″	15°23′13″	3	6.88	3.64
WS-6	2006	ITC	10	27°51′07″	15°24′29″	7	6.02	3.59
WS-7	2006	ITC	20	27°51′07″	15°24′29	7	6.95	3.62
WS-8	2006	ITC	13	28°09′14″	15°31′44″	15	3.53	2.32
WS-9	2006	ITC	15	28°09′47″	15°42′25″	57	8.21	3.97
WS-10	2006	ITC	20	28°10′12″	15°38′20″	30	5.86	3.18
WS-11	2006	ITC	10	28°10′12″	15°38′20″	30	5.60	3.09
WS-12	2006	ITC	15	28°00′00″	15°48′47″	6	5.15	2.94
WS-13	2006	ITC	20	27°57′36″	15°33′36″	1905	6.61	4.37
WS-14	2006	ITC	25	27°51′47″	15°25′12″	37	6.08	3.39
WS-15	2002	ITC	10	27°53′24″	15°23′42″	3	6.36	3.38
WS-16	2006	ITC	10	28°06′36″	16°29′06″	79	5.35	3.66
WS-17	2002	ITC	10	28°25′12″	16°21′18″	732	6.26	5.26
WS-18	2002	AEMET	10	28°02′35″	16°34′16″	51	5.53	3.26
WS-18	2006	AEMET	10	28°02′35″	16°34′16″	51	5.17	2.91
WS-19	2002	AEMET	10	28°27′10″	13°51′54″	24	5.88	2.66
WS-19	2006	AEMET	10	28°27′10″	13°51′54″	24	5.78	2.51
WS-20	2002	AEMET	10	28°36′47″	17°45′36″	85	4.82	2.30
WS-20	2006	AEMET	10	28°36′47″	17°45′36″	85	4.45	2.23
WS-21	2002	AEMET	10	27°48′50″	17°53′10″	30	5.84	3.16
WS-22	2002	AEMET	10	28°57′07″	13°36′00″	10	5.99	3.14
WS-22	2006	AEMET	10	28°57′07″	13°36′00″	10	5.55	2.83

Table 2
Linear correlation coefficients between the wind speeds of the different anemometer weather stations. Year 2002.

	WS-1	WS-2	WS-3	WS-4	WS-15	WS-17	WS-18	WS-19	WS-20	WS-21	WS-22
WS-1	1.000	0.939	0.712	0.705	0.952	0.417	0.506	0.690	0.601	0.492	0.641
WS-2	0.939	1.000	0.676	0.670	0.913	0.467	0.460	0.706	0.630	0.504	0.632
WS-3	0.712	0.676	1.000	0.999	0.741	0.241	0.604	0.588	0.506	0.421	0.564
WS-4	0.705	0.670	0.999	1.000	0.735	0.246	0.601	0.587	0.506	0.423	0.563
WS-15	0.952	0.913	0.741	0.735	1.000	0.412	0.546	0.696	0.608	0.493	0.650
WS-17	0.417	0.467	0.241	0.246	0.412	1.000	0.167	0.464	0.506	0.415	0.404
WS-18	0.506	0.460	0.604	0.601	0.546	0.167	1.000	0.512	0.393	0.300	0.511
WS-19	0.690	0.706	0.588	0.587	0.696	0.464	0.512	1.000	0.562	0.481	0.679
WS-20	0.601	0.630	0.506	0.506	0.608	0.506	0.393	0.562	1.000	0.476	0.496
WS-21	0.492	0.504	0.421	0.423	0.493	0.415	0.300	0.481	0.476	1.000	0.506
WS-22	0.641	0.632	0.564	0.563	0.650	0.404	0.511	0.679	0.496	0.506	1.000

Table 3
Linear correlation coefficients between the wind speeds of the different anemometer weather stations. Year 2006.

	WS-1	WS-5	WS-6	WS-7	WS-8	WS-9	WS-10	WS-11	WS-12	WS-13	WS-14	WS-16	WS-18	WS-19	WS-20	WS-22
WS-1	1.000	0.903	0.892	0.887	0.200	0.625	0.414	0.407	0.728	0.091	0.885	0.662	0.459	0.647	0.575	0.687
WS-5	0.903	1.000	0.902	0.896	0.197	0.712	0.518	0.504	0.717	0.109	0.902	0.671	0.450	0.633	0.559	0.662
WS-6	0.892	0.902	1.000	0.995	0.145	0.687	0.479	0.470	0.695	0.040	0.963	0.710	0.515	0.614	0.528	0.668
WS-7	0.887	0.896	0.995	1.000	0.146	0.691	0.482	0.475	0.692	0.053	0.962	0.712	0.521	0.615	0.544	0.672
WS-8	0.200	0.197	0.145	0.146	1.000	0.350	0.514	0.521	0.161	0.306	0.159	0.281	0.275	0.430	0.369	0.265
WS-9	0.625	0.712	0.687	0.691	0.350	1.000	0.809	0.802	0.528	0.272	0.688	0.699	0.546	0.553	0.510	0.515
WS-10	0.414	0.518	0.479	0.482	0.514	0.809	1.000	0.996	0.346	0.225	0.493	0.643	0.559	0.498	0.414	0.402
WS-11	0.407	0.504	0.470	0.475	0.521	0.802	0.996	1.000	0.346	0.226	0.485	0.639	0.562	0.497	0.413	0.400
WS-12	0.728	0.717	0.695	0.692	0.161	0.528	0.346	0.346	1.000	0.151	0.701	0.523	0.353	0.510	0.479	0.570
WS-13	0.091	0.109	0.040	0.053	0.306	0.272	0.225	0.226	0.151	1.000	0.069	0.145	0.059	0.215	0.301	0.131
WS-14	0.885	0.902	0.963	0.962	0.159	0.688	0.493	0.485	0.701	0.069	1.000	0.721	0.523	0.624	0.560	0.667
WS-16	0.662	0.671	0.710	0.712	0.281	0.699	0.643	0.639	0.523	0.145	0.721	1.000	0.780	0.590	0.503	0.579
WS-18	0.459	0.450	0.515	0.521	0.275	0.546	0.559	0.562	0.353	0.059	0.523	0.780	1.000	0.481	0.372	0.455
WS-19	0.647	0.633	0.614	0.615	0.430	0.553	0.498	0.497	0.510	0.215	0.624	0.590	0.481	1.000	0.504	0.651
WS-20	0.575	0.559	0.528	0.544	0.369	0.510	0.414	0.413	0.479	0.301	0.560	0.503	0.372	0.504	1.000	0.503
WS-22	0.687	0.662	0.668	0.672	0.265	0.515	0.402	0.400	0.570	0.131	0.667	0.579	0.455	0.651	0.503	1.000

range of coefficients of correlation varies between 0.04 and 0.999. The lowest value was obtained between stations WS-13 and WS-6 on the island of Gran Canaria (Fig. 2). The first of these is located in the centre of the island 1905 m above sea level, while the second is found on the eastern coast of the island at 7 m above sea level. The highest correlation coefficients were observed between the data measured by the recording devices at the following stations: WS-3 and WS-4; WS-6 and WS-7; and WS-10 and WS-11. Each of these is a pair of recording devices installed on the same supporting pole, but at different heights above ground level.

The prevailing wind regime in the Canary Archipelago is the trade wind regime, which blows for a high percentage of the time in a north-easterly direction. Fig. 3a shows, by way of representative example of the prevailing wind regime in the Canary Islands, the wind rose for station WS-5. There are, however, some stations which as a result of their orographic conditions have wind roses with other prevailing directions. Fig. 3b shows the wind rose for WS-12 as a representative example of these stations.

The wind turbine models used in this study [20,21] were chosen so as to correspond to different ranges of rated power. The power-wind speed characteristic curves of these turbines are shown in Fig. 4.

4. Methodology

The aim of the methodology proposed in this paper is to analyse the influence that the number of reference stations, the degree of correlation between the wind speeds of the reference stations and the candidate station, the wind direction and the manner in which the direction signal is introduced into the input layer of the neural

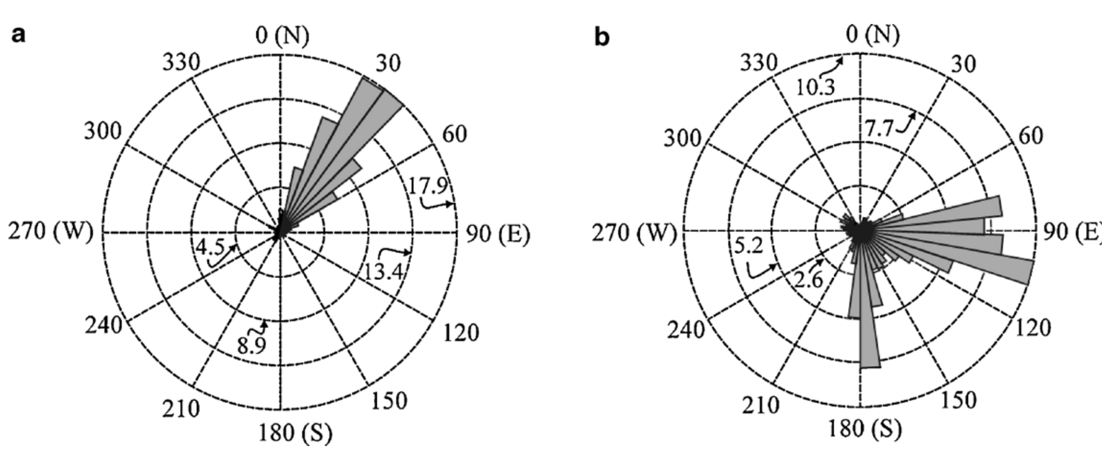


Fig. 3. Wind roses of the weather stations in 2006: (a) WS-5, (b) WS-12.

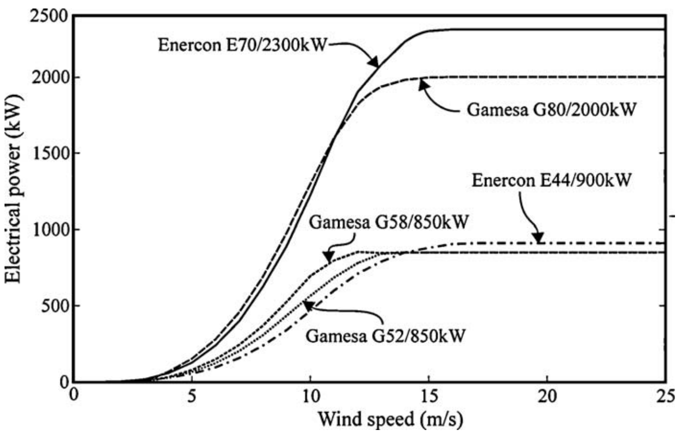


Fig. 4. Power-wind speed characteristic curves of the wind turbines used in the study.

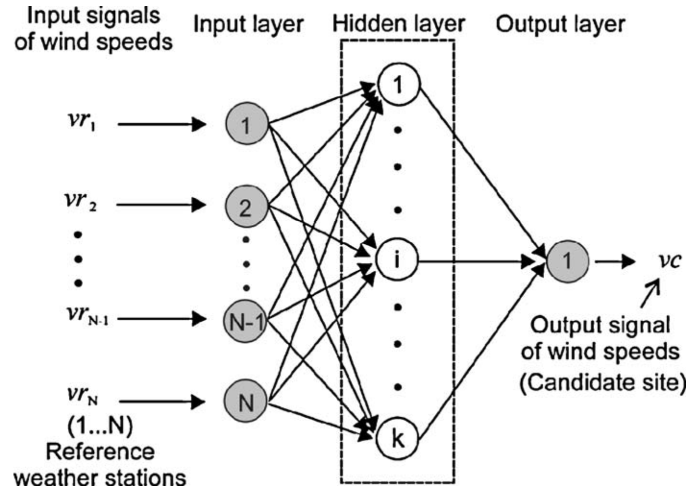


Fig. 5. Schematic diagram of an ANN with N wind speed input signals of N reference stations and one wind speed output signal of the target station.

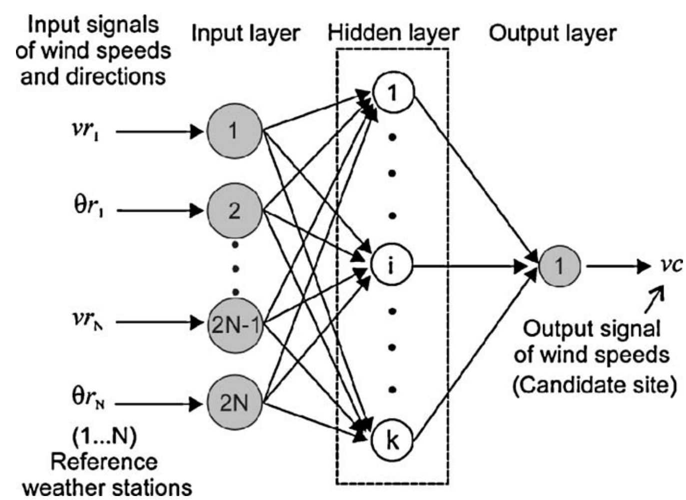


Fig. 6. Schematic diagram of an ANN with $2N$ wind speed and angular wind direction input signals of N reference stations and one wind speed output signal of the target station.

networks has on the estimation of three variables of the candidate station. These variables are: (a) the mean hourly wind speeds, (b) the mean hourly wind power densities and (c) the hourly electrical power output of the wind turbines to be installed at the candidate site.

Measurement of the different influences was performed via the analysis of the errors and relationships obtained in the evaluations carried out with the set of data not used in the training or validation of the networks. These errors and relationships were evaluated with two metrics. These were the mean absolute relative error (MARE), Eq. (1), and the correlation coefficient (CC), Eq. (2), which have also been used by other authors [6,9].

$$\mathrm{MARE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|T_i - O_i\right|}{T_i} \tag{1}$$

$$\mathrm{CC} = \frac{\sum_{i=1}^{n}(T_i - \bar{T})(O_i - \bar{O})}{\sqrt{\left[\sum_{i=1}^{n}(T_i - \bar{T})^2\right]\left[\sum_{i=1}^{n}(O_i - \bar{O})^2\right]}} \tag{2}$$

where n is the number of data of the test subset; O_i are the data estimated with the model; $\bar{O}$ is the mean of the values O_i; T_i are the measured or real data and $\bar{T}$ is the mean of the values T_i.

The methodological process followed was as described below:

(A) Determination of the errors committed when the architecture of the neural network used has only one input signal and only one output signal. That is, in Fig. 5, $N = 1$. The input signal is the series of wind speeds of a reference station while the output signal is the series of wind speeds of the target station.

At this point all the stations can act as reference and as target station. Bearing in mind the number of available stations (Table 1), the number of possible combinations is 55 for the year 2002 and 120 for the year 2006. So, the number of cases analysed was 175.

(B) Determination of the errors committed when the architecture of the neural network used has two input signals and only one output signal. That is, in Fig. 6, $N = 1$. The input signals are the series of wind speeds and directions of a reference station and the output signal the series of wind speeds of the target station. In these models the wind direction signal is introduced in angular magnitude (0–360°) and the angle corresponding to the northerly direction is taken as angle 0°. As in the previous case, all the stations can act as reference and as target station. So, the number of cases analysed here is also 175.

(C) Determination of the errors committed when the architecture of the neural network used has two input signals and two output signals. That is, in Fig. 7a, $N = 1$. The input signals are the two series of Cartesian components (vrx,vry) of the wind velocity of the reference station and the output signals are the two series of Cartesian components (vcx,vcy) of the wind velocity of the target station. The wind direction is measured in a clockwise direction[1] using the North as starting point (the y axis in Fig. 7b). The modules vc_i of the series of output velocities of the 175 cases analysed are obtained from Eq. (3).

$$vc_i = \sqrt{vcx_i^2 + vcy_i^2} \tag{3}$$

(D) Determination of the errors committed when the architecture of the neural network used has $2N$ ($N > 1$) input signals and only one output signal (Fig. 6). The $2N$ input signals are the series of wind speeds and directions of N reference stations and the output signal the series of wind speeds of the target station.

(E) Determination of the errors committed when the architecture of the neural network used has N ($N > 1$) input signals and only

[1] In this paper the angle corresponding to the northerly direction is taken as angle 0°.

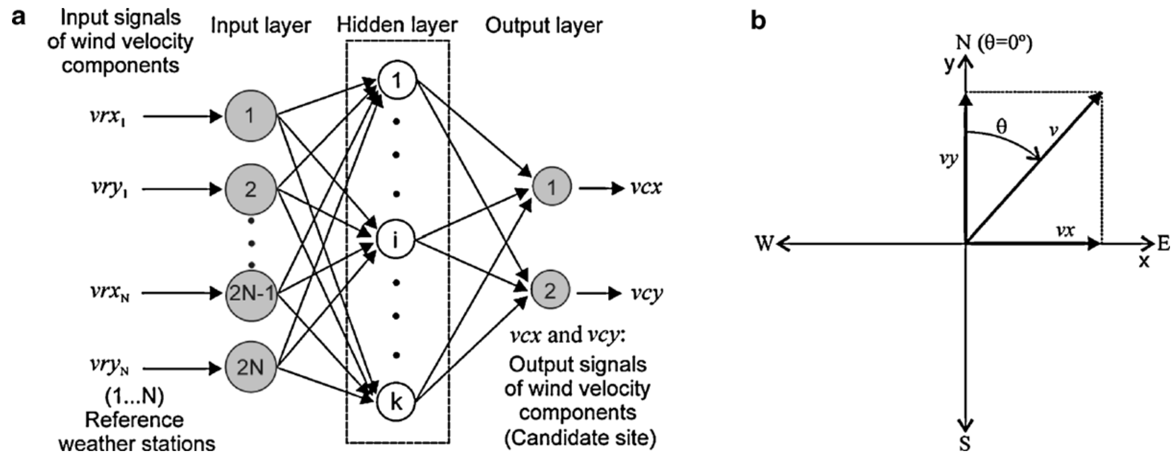


Fig. 7. (a) Schematic diagram of an ANN with 2N input signals of the Cartesian components of the wind velocity of N reference stations and two output signals of the Cartesian components of the wind velocity of the target station. (b) Cartesian decomposition of the wind velocity.

one output signal (Fig. 5). The N input signals are the series of wind speeds of N reference stations and the output signal the series of wind speeds of the target station.

(F) Determination of the errors committed when the architecture of the neural network used has 2N input signals and two output signals (Fig. 7a). The 2N input signals are the series of Cartesian components of the wind velocities of the N ($N>1$) reference stations and the output signals are the two series of Cartesian components (vcx,vcy) of the wind velocity of the target station.

(G) Determination, for each of the previous sections, of the errors committed in the estimation of the wind power density, P_i, which is determined through Eq. (4) [25].

$$P_i = \frac{1}{2}\rho_i v_i^3 \tag{4}$$

In Eq. (4), if ρ_i (air density) is expressed in kg m^{-3} and v_i (wind speed) is expressed in m s^{-1}, P_i is obtained in W m^{-2}. Air density is a continuous variable that depends on pressure, temperature and humidity [25]. However, in the scientific literature, it is assumed that the air density is constant. A typical value used in all the literature consulted is $\bar{\rho} = 1.225$ kg m^{-3}, corresponding to standard conditions (sea level, 15 °C) [25].

(H) Determination, for each of the previous sections, of the errors committed in estimation of the mean power output of the five wind turbines selected (Fig. 4). The manufacturers usually supply the power curves of their wind turbines in a discrete form with M nodes (Pwt_j,V_j). The electrical power obtainable between two nodes "j" and "$j+1$" of the power curve can be calculated approximately [26]. One of the possible means of approximation consists of assuming that the variation between two nodes of the power wind speed curve is linear, Eq. (5)

$$PWT(v_i) = \begin{cases} \dfrac{Pwt_{j+1} - Pwt_j}{V_{j+1} - V_j}(v_i - V_j) + Pwt_j \\ 0; \quad V_{cut-out} < v_i < V_{cut-in} \end{cases} \tag{5}$$

where $V_{cut\text{-}in}$ and $V_{cut\text{-}out}$ are the cut-in and cut-out speeds, respectively, of the wind turbine. It should be mentioned that the wind speeds used in Eq. (5) are those estimated for the target station at the height indicated in Table 1. This has been done in this way so as not to introduce the errors that are generated with extrapolation models of wind speed with height.

(I) Comparative analysis of the results obtained in the previous sections.

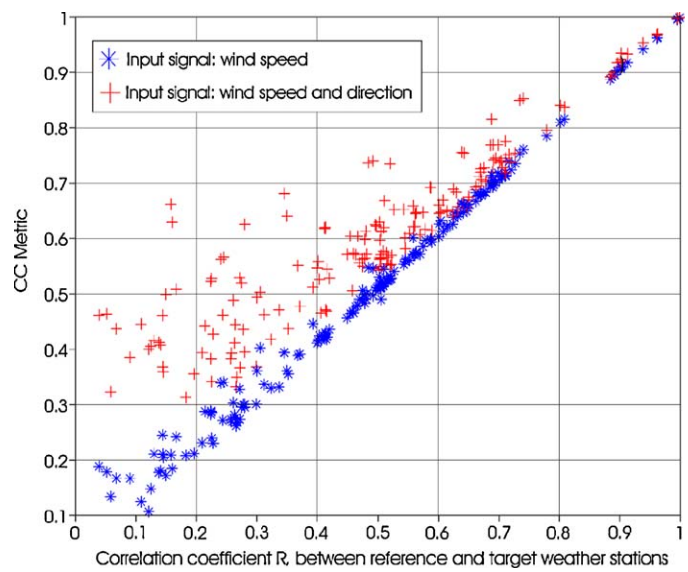


Fig. 8. CC metrics of cases A and B as a function of the coefficients of correlation R existing between the wind speeds of the reference and target stations.

5. Analysis of results

Fig. 8 shows a comparison of the results obtained for cases A and B of the previous section. This comparison is performed by representing on the y-axis the CC metrics estimated with Eq. (2), and on the x-axis the existing coefficients of correlation R between the wind speeds measured at the reference and target stations (Tables 2 and 3). The same comparison is undertaken in Fig. 9, but this time employing the MARE metric, Eq. (1), instead of the CC metric.

It can be seen in both figures that when the wind direction in angular magnitude is used as an input signal of the neural networks, the metrics obtained in all the analysed cases give better results than those obtained when this direction signal is ignored. However, when the coefficients of correlation R are higher than 0.98 the influence of the direction signal is of low significance. Likewise, it can be seen from Figs. 8 and 9 that the favourable influence of consideration of the direction signal is most notable when the coefficients of correlation R have values lower than 0.6. In these cases, the metrics obtained can display improvements (of up to 100% in the case of CC and 20% in the case of MARE) over the results obtained when ignoring the direction signal. In addition, for values of R ranging

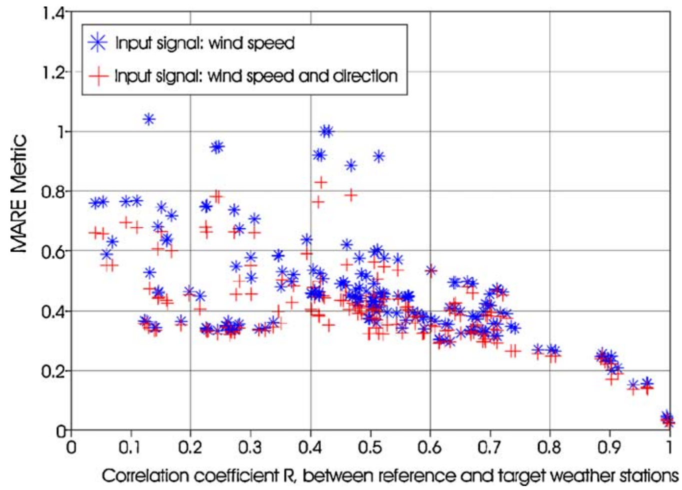


Fig. 9. MARE metrics of cases A and B as a function of the coefficients of correlation *R* existing between the wind speeds of the reference and target stations.

between 0.6 and 0.98 inclusion of direction as an input signal of the neural networks also has a favourable influence on estimation of the wind speeds of the target station. In this range the CC showed improvements of up to 15% and the MARE of up to 20%.

Figs. 10 and 11 show the results obtained in the comparison of cases B and C of the methodological procedure that was followed. That is, a comparison was undertaken of the manner of introducing the direction signal in the neural network models. The conclusion was drawn that for values of *R* higher than 0.9 there are no significant differences between the results obtained when the direction signal is introduced explicitly in angular magnitude and when it is introduced implicitly by decomposing the horizontal wind velocity into its Cartesian components. However, for values of *R* ranging between 0.6 and 0.9 the models that decompose the velocity into its Cartesian components display in most cases worse metrics. For values of R lower than 0.6 no predominant effect was observed of either of the two ways of introducing the direction signal.

Figs. 12 and 13 represent, by way of example, the results obtained with some ANNs of case D as indicated in the methodology, when $N = 2$. That is, when dealing with ANNs with four input signals and one output signal. These figures aim to show the results obtained when the input signals of a second reference station, which has a value of *R* with respect to the candidate station lower than that of the first reference station, are added to the input sig-

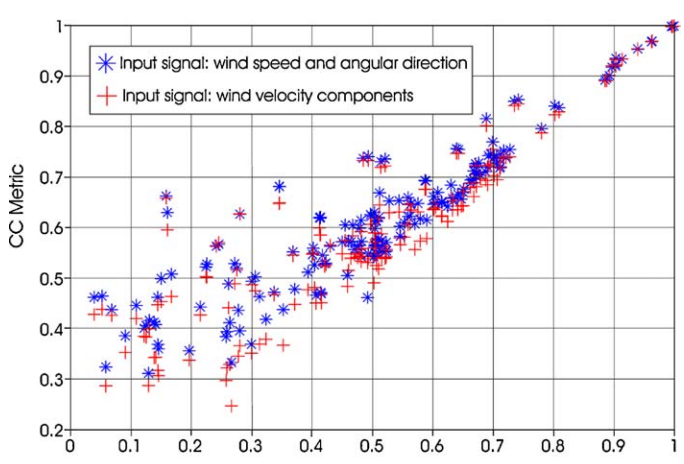


Fig. 10. CC metrics of cases B and C as a function of the coefficients of correlation *R* existing between the wind speeds of the reference and target stations.

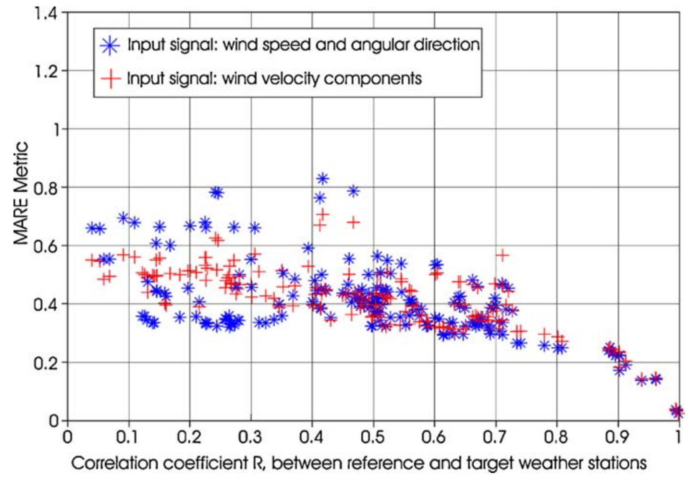


Fig. 11. MARE metrics of cases B and C as a function of the coefficients of correlation *R* existing between the wind speeds of the reference and target stations.

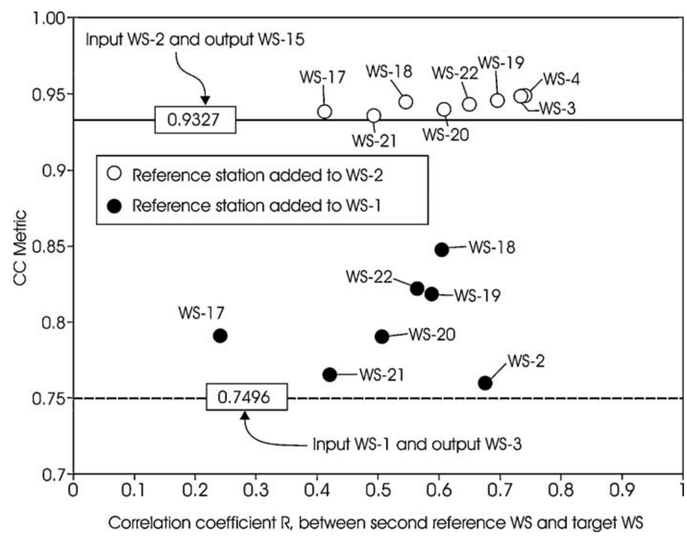


Fig. 12. Comparison of the CC metric of two models of case A with the corresponding metric of various models of case D with $N = 2$. The D models were generated adding to the case A models a second reference station which has an *R* value with respect to the candidate station that is lower than that held by the first station.

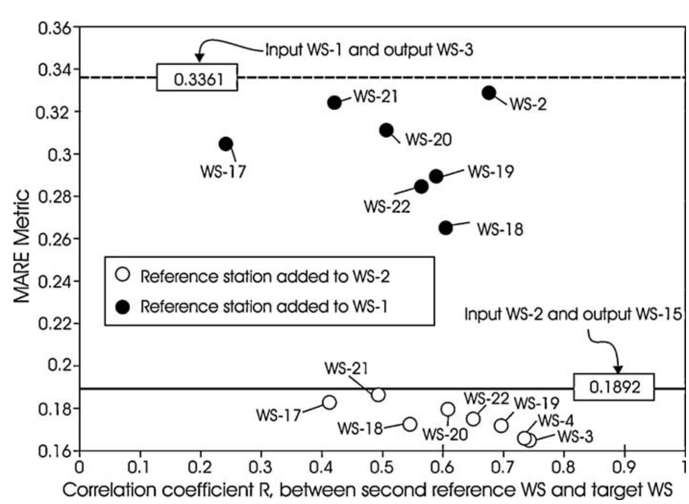


Fig. 13. Comparison of the MARE metric of two models of case A with the corresponding metric of various models of case D with $N = 2$. The D models were generated adding to the case A models a second reference station which has an *R* value with respect to the candidate station that is lower than that held by the first station.

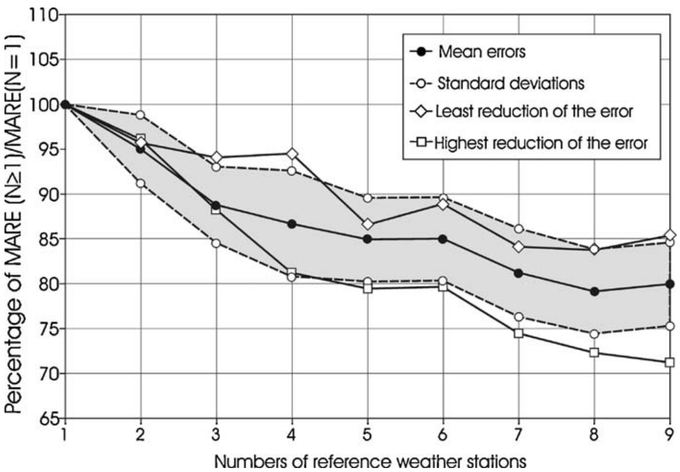


Fig. 14. Percentage of the ratio between the MARE obtained when *N* reference stations are used (case E) and the MARE obtained when only one reference station is used (case A), against the number of reference stations *N* used in estimation of the wind speed of a target station.

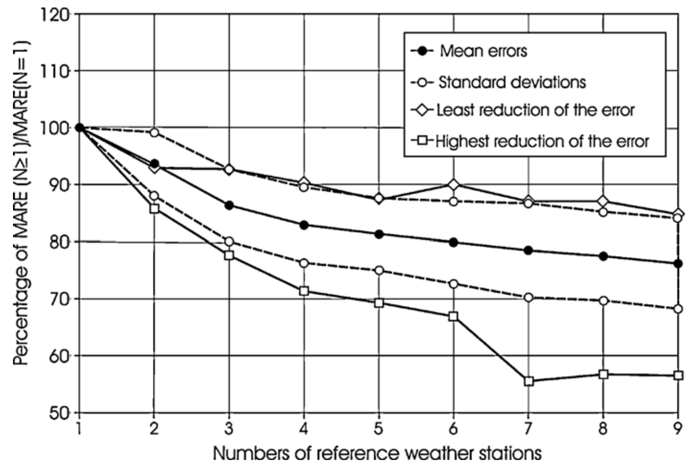


Fig. 15. Percentage of the ratio between the MARE obtained when *N* reference stations are used (case D) and the MARE obtained when only one reference station is used (case B), against the number of reference stations *N* used in estimation of the wind speed of a target station.

nals of case B (as indicated in the methodology). The purpose here is to analyse the effect of the *R* value of the second reference station on the magnitude of the error metrics employed. Fig. 12 shows the CC metric against *R* and Fig. 13 the MARE metric against *R*. Figs. 12 and 13 use two of the results represented in Figs. 8 and 9, respectively. Specifically, one of the results is that obtained by the ANN whose input signals come from the WS-1, while the output signal is the estimated wind speed of station WS-3. The other result is that obtained by the ANN with reference station WS-2 and target station WS-15.

Stations WS-1 and WS-3 gave a value of *R* in the year 2002 of 0.712, while the CC and MARE metrics obtained in case B were 0.7496 and 0.3361, respectively. These metrics are represented with dashed lines in Figs. 12 and 13. Stations WS-2 and WS-15 give an *R* value of 0.913 in the year 2002, while the CC and MARE metrics obtained in case B are 0.9327 and 0.1892, respectively. These values are also represented with a horizontal line, this time continuous, in Figs. 12 and 13.

It can be observed in Figs. 12 and 13 that the metrics improve when a second reference station is added, independently of the magnitudes of the coefficients of correlation *R* that exist between the first reference station and the target station. It should however be pointed out that the percentage improvement tends to decrease as *R* increases, with the improvement being nullified when *R* reaches the unit value. It can be deduced from Figs. 12 and 13 that with the addition of WS-18 to WS-1 the CC increases by 13.02% (from 0.7496 to 0.8472) and the MARE decreases by 21.15% (from 0.3361 to 0.265). However, with the addition of WS-3 to WS-2 the CC only increases by 1.4% (from 0.9327 to 0.9457) and the MARE decreases by 12.7% (from 0.189 to 0.165). It is also observed that the percentage improvement of the metrics does not display a dependency on the coefficient of correlation *R* that the second reference station has with the target station. The results that are shown by way of example in Figs. 12 and 13 were likewise observed in all the combinations carried out with the remaining stations.

Fig. 14 shows the results obtained in case E when *N* varies. Shown on the *y*-axis of Fig. 14 is the percentage of the ratio between the MARE obtained when *N* reference stations are used and the MARE obtained when only one reference station is employed. On the *x*-axis is the number of reference stations N used in the estimation of the wind speed at a target station. Fig. 14 was drawn using the ANNs analysed in case B to which were added the reference stations that had a value of *R* with respect to the candidate station lower than that of the first reference station. The mean values and the standard deviations of the percentages of variation of the MARE metric are shown in this figure. Also shown are the extreme situations of all the ANNs generated. That is, the situations that displayed highest and lowest variation of the error. The decreasing tendency can be observed in Fig. 14 of the variations of the MARE metric as the number of reference stations increases. It should be pointed out, however, that not all the added reference stations improve the results. So, the trend curves can display saw-tooth forms. It should also be mentioned that from the tests conducted with all the possible combinations of stations it was observed that elimination of reference stations which increase the error does not affect obtainment of the minimum error. So, in order to detect the minimum error when various reference stations are available for estimation of the wind speed of a target station it is recommended that a graph such as that shown in Fig. 14 be made. The conclusions extracted from Fig. 14 can be extrapolated to case F. Fig. 15 shows a similar analysis to that of Fig. 14, but for case D when *N* varies using as input signals the wind direction in angular magnitude as well as the wind speed. The conclusions extracted from Fig. 15 are similar to those for Fig. 14. However, if the mean trends of the variation of the error of both figures are compared, it is seen that the error decrease percentages are higher when the wind directions in angu-

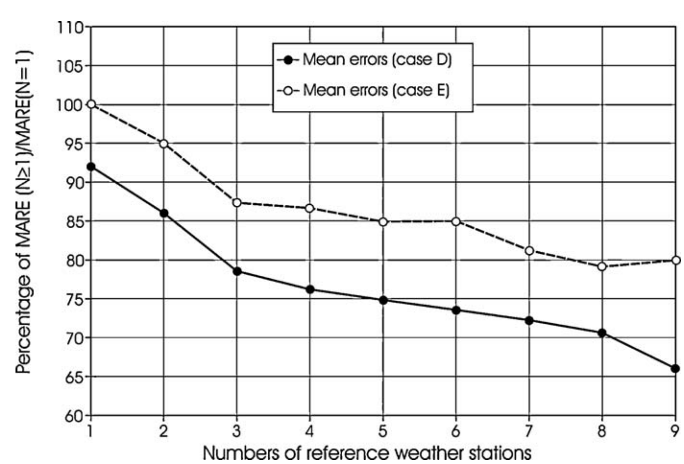


Fig. 16. Percentage of the ratio between the MARE obtained when *N* reference stations are used (case D and case E) and the MARE obtained when only one reference station is used (case A), against the number of reference stations *N* used in estimation of the wind speed of a target station.

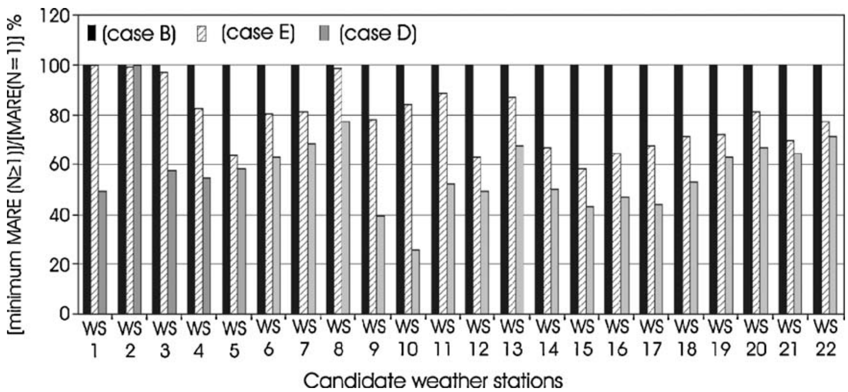


Fig. 17. Percentage of the ratio between the minimum MARE obtained when analysing the mean wind power densities in cases D and E and the MARE obtained when only one reference station is used (case B) against the target station considered.

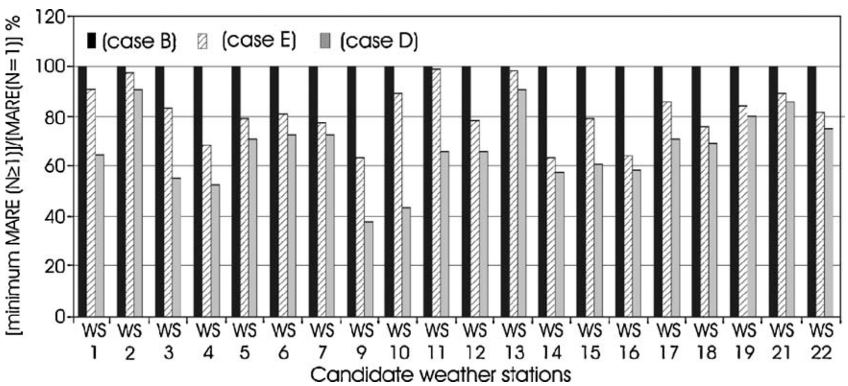


Fig. 18. Percentage of the ratio between the minimum MARE obtained when analysing the mean power output of the wind turbines in cases D and E and the MARE obtained when only one reference station is used (case B) against the target station considered.

lar magnitude are used as input signals of the ANNs. Shown along the *y*-axis of Fig. 16 is the percentage of the ratio between the MARE obtained in cases E and D and the MARE obtained when only one reference station is used with input signals of wind speed alone. It can be seen for this figure that if wind speed and wind angular direction input signals are used then the number of reference stations required to achieve a particular magnitude of the metric is lower than the number of stations required if the input signals are wind speed alone.

The various stations analysed in this paper are represented along the *x*-axis of Fig. 17. The minimum MARE obtained when *N* varies in estimation of the wind power density and the MARE obtained in that estimation when only one reference station is used (case B) are compared along the *x*-axis of this figure. The value 100% was assigned to the error of the latter case. This enables comparison of the errors committed in wind power density estimation when only one reference station is used with the errors produced in the following two hypotheses: (1) when the optimum number of reference stations is used, but employing only wind speed signals and (2) when the optimum number of reference stations is used, but wind speed and direction (in angular magnitude) signals are used. It can be observed in Fig. 17 that, for all the stations analysed, the minimum error is achieved when various reference stations are employed and the ANNs are fed with wind speed and wind direction signals.

Given the existing power relationship between wind power density and wind speed, small reductions in the estimation errors of the latter can give rise to significant reductions in the estimation of the former. It can be observed in Fig. 17 that in the estimation of the wind power density of the WS-10, in which three reference stations were used, a 75% reduction of the MARE was achieved with respect to the MARE of case B. Fig. 18 provides a similar analysis to that of Fig. 17, but for estimation of the power output of the five turbines under consideration (Fig. 4). When making Fig. 18 only those candidate stations whose mean annual wind speed was equal to or greater than 5 m/s were considered. It should be mentioned that very significant differences were not detected between the errors generated by the different wind turbines used. The standard deviations of the errors were never higher than 5% of the mean error. The conclusions drawn after analysis of the results are similar to those expressed for the case of estimation of the wind power density. In the case of the WS-9, using seven reference stations, a 62% reduction of the MARE was achieved with respect to the case B MARE.

6. Conclusions

A presentation is made in this paper of the results of a study carried out in the Canary Archipelago to analyse the influence that the characteristics of the input signals of ANN models have on estimation of three variables of the candidate station. These variables are: (a) the mean hourly wind speeds, (b) the mean hourly wind power densities and (c) the mean hourly power output of five wind turbines that could be installed at the candidate site.

After application of the methodology as explained in this paper to a set of weather stations installed in the Canary Islands a series of conclusions has been drawn that may be helpful in the use of ANN models to estimate the wind energy potential of a site for which no long-term wind records are available. It is concluded from the study that: (a) when the wind direction in angular magnitude is used as an input signal of the ANNs, the metrics obtained give better results than those obtained when the wind direction signal

is omitted. The decrease of the errors is noticeable fundamentally when the *R*s existing between the candidate station and the reference stations are not high; (b) for values of *R* between 0.6 and 0.9 the models that decompose the wind speed into its Cartesian components show, in most cases, worse metrics than the models which use wind direction in angular magnitude as an input signal; (c) the estimation errors tend to decrease as the number of reference stations used increases. However, not all the reference stations added lead to obtainment of the absolute optimum. So, it is proposed that graphs such as that shown in Fig. 15 be made for the purpose of determining the reference stations which provide the minimum estimation error; (d) if wind speed and angular wind direction input signals are used then the number of reference stations required to achieve a particular magnitude of the metric is lower than the number of stations required if only wind speed input signals are used; (e) by using multiple reference stations with wind speed and angular wind direction input signals it was possible to reduce the error in estimation of wind power density and the electrical power of some stations, with respect to case B, by 75% and 62%.

References

[1] Barros VR, Estevan EA. On the evaluation of wind power from short wind records. Journal of Climate and Applied Meteorology 1983;22:1116–23.

[2] Justus CG, Mani K, Mikhail AS. Interannual and month-to-month variations of wind speed. Journal of Applied Meteorology 1979;18:913–20.

[3] Daniels PA, Schroede TA. Siting large wind turbines in Hawaii. Wind Engineering 1988;12:302–30.

[4] Carta JA, González J. Self-sufficient energy supply for isolated communities: wind-diesel systems in the Canary Islands. The Energy Journal 2001;22:115–45.

[5] García-Rojo R. Algorithm for the estimation of the long-term wind climate at a meteorological mast using a joint probabilistic approach. Wind Engineering 2004;28:213–24.

[6] Hiester TR, Pennell WT. The siting handbook for large wind energy systems. 2nd ed. New York: WindBook; 1981.

[7] Witten IH, Frank E. Data mining. Practical machine learning tools and techniques. 2nd ed. San Francisco: Elsevier; 2005.

[8] Rogers AL, Rogers JW, Manwell JF. Comparison of the performance of four Measure-Correlate-Predict algorithms. Journal of Wind Engineering and Industrial Aerodynamics 2005;93:243–64.

[9] Woods JC, Watson SJ. A new matrix method of predicting long-term wind roses with MCP. Journal of Wind Engineering and Industrial Aerodynamics 1997;66:85–94.

[10] Clive JM. Non-linearity in MCP with Weibull distributed wind speeds. Wind Engineering 2008;32:319–24.

[11] Oztopal A. Artificial Neural Network approach to spatial estimation of wind velocity. Energy Conversion and Management 2006;47:395–406.

[12] Bilgili M, Yasar SBA. Application of artificial neural networks for the wind speed prediction of target station using reference stations data. Renewable Energy 2007;32:2350–60.

[13] Mabel C, Fernandez E. Analysis of wind power generation and prediction using ANN: a case study. Renewable Energy 2008;33:986–92.

[14] Monfared M, Rastegar H, Kojabadi HM. A new strategy for wind speed forecasting using artificial intelligent methods. Renewable Energy 2009;34:845–8.

[15] Fadare DA. The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. Applied Energy 2010;87:934–42.

[16] Lopez P, Velo R, Maseda F. Effect of direction on wind speed estimation in complex terrain using neural networks. Renewable Energy 2008;33:2266–72.

[17] Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews 2001;5:373–401.

[18] Carta JA, Velázquez S, Matías JM. Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site. Energy Conversion & Management 2011;52:1137–49.

[19] Draper NR, Smith H. Applied regression analysis. 1st ed. New York: John Wiley & Sons, Inc.; 1998.

[20] Enercon GmbH. Enercon booklets. Enercon wind turbines. Product overview. Available from: www.Enercon.de.

[21] Gamesa. Gamesa products. Wind turbines. Wind turbines catalogue. Available from: www.gamesa.com/en.

[22] Principe JC, Euliano NR, Lefebvre WC. Neural and adaptive systems. Fundamentals through simulations. 1st ed. New York: John Wiley & Sons, Inc.; 2000.

[23] Masters T. Practical neural network recipes in C++. 1st ed. California: Morgan Kaufmann Publishers; 1993.

[24] Demuth H, Beale M. Neural network toolbox user's guide. Natick, MA 01760-2098: The Math Works, Inc.; 2003.

[25] Carta JA, Mentado D. A continuous bivariate model for wind power density and wind turbine energy output estimations. Energy Conversion & Management 2007;48:420–32.

[26] Carta JA, Ramírez P, Velázquez S. Influence of the level of fit of a density probability function to wind-speed data on the WECS mean power output estimation. Energy Conversion & Management 2008;49:2647–55.